Music performance research at the millennium

ALF GABRIELSSON
UPPSALA UNIVERSITY, SWEDEN

ABSTRACT Empirical research on music performance has increased considerably during recent decades. This article updates the review of the research up to 1995 published by the current author in 1999. Covering about 200 papers from 1995 up to 2002, this article confirms the impression that music performance research is in a very active stage. As in the previous review, the majority of papers are on measurement of performance, but there is a rapidly increasing number of contributions concerning models of performance, performance planning and practice. Although fewer in number, there are also many new contributions within each of the remaining areas of performance research analysed in this review.

KEYWORDS: music psychology, performance, review

Music performance has been the subject of many treatises by musicians and music theorists in the past. To mention but a few, observations and advice on proper performance were provided in classical works such as Mattheson (1739), Quantz (1752), C.P.E. Bach (1753, 1762) and L. Mozart (1756) — works that are still frequently cited. Contemporary works that provide much relevant historical information on performance are Hudson (1994), the thesis by Humair (1999) and articles by Kopiez (1996), Lawson (2002), Lehmann and Ericsson (1998b), and Walls (2002).

It was not until around 1900 that empirical studies of music performance began to appear in journals of the new science called psychology. At the very beginning, the focus was on physical measurements of music performances, especially regarding timing in performance. After some early research in Europe and the USA, this work reached a peak during the 1920s and 1930s in extensive studies of performance on the piano, the violin and in singing conducted by researchers at Iowa University led by Carl Emil Seashore. This
research was reported in several volumes that are little read today; however, a survey was published in Seashore’s textbook on music psychology (1938). Seashore’s retirement and the Second World War brought an end to these activities and performance research remained dormant until a revival in the 1960s. A pioneer in this revival was my teacher in musicology, Professor Ingmar Bengtsson at Uppsala University, who guided me in research of rhythm and performance. Measurements of several types of music performances – orchestra, organ, singing, jazz and Swedish folk music – were described in internal reports (e.g. Bengtsson, 1967), and physicists in Uppsala constructed advanced new equipment for the analysis of monophonic music performance (Tove et al., 1966). In a comprehensive survey of this early work in Uppsala (Bengtsson et al., 1969), Ingmar put forth a hypothesis concerning systematic variations (SYVAR) in timing, which was exemplified by measurements of Viennese waltzes and Swedish folk music. The paper also included an extensive summary of psychological rhythm research that was my main contribution. We later regretted that this research was only published in Swedish; however, the main themes regarding the empirical topics also appeared in later papers during the 1970s (Bengtsson, 1974; Bengtsson and Gabrielsson, 1977).

Since about 1975, there has been an increasing amount of performance research. In working on a review of empirical performance research up to the mid-1990s (Gabrielsson, 1999a), I found around 500 works, many more than I had expected. I organized this material in a type of chronological order, going from the initial planning of performance through many different aspects of performance itself (the most comprehensive part of the review) to a final evaluation of performance (see Table 1).

The approximate number of papers within each of the main topics in Table 1 is shown in Figure 1. The largest number of papers concerned measurement of performance; in addition, many of the papers dealing with motor processes and models of performance also referred to measurements of various kinds. Papers on psychological and social factors were the next highest number; about half of the papers in this category concerned performance anxiety. The remaining topics had so far attracted relatively less interest, especially matters concerning improvisation, feedback and evaluation. (As there is some overlap of different categories, and as many papers pertained to phenomena in two or more of the categories, the total number of references in the figure is larger than the approximate 500 papers reviewed.) Practically all the papers dealt with performance of western tonal music, mostly art (‘classical’) music.

For comparison – and for examples of other organizational structures and focus of topics – the reader should consult two other recent reviews of performance research (Kopiez, 1996; Palmer, 1997).

The first version of my review was ready in 1992. Since the publication of the book in which it should have appeared was delayed, I revised my manuscript
in 1995 adding about 100 papers that had appeared during the intervening three years (1992–5). At the time this article was written (2002), there are of course yet more papers. Browsing through the most prominent journals in music psychology from 1995 onwards and consulting some databases on
The purpose of this article is to provide an update of my 1999 review (Gabrielsson, 1999a, abbreviated to G99 from now on). To keep this update within reasonable limits, I decided to include only papers published in regular journals or books. Papers in conference proceedings and unpublished doctoral theses are excluded, except for a few cases. As in G99, I only briefly refer to matters directly related to music teaching and learning; for these, I refer readers to the recent volumes edited by Jorgensen and Lehmann (1997), Parncutt and McPherson (2002), and Rink (2002). I make no claim that this review is exhaustive, and I thus apologize to colleagues whose paper(s) may be missing or only briefly described.

The general background for this article is, of course, the review in G99. In the present article, I make a few statements at the beginning of each section which link with the corresponding section in G99. However, for a more
complete picture, it is necessary to consult G99. I have organized the material according to the same principles as in G99 (see Table 1) but the order of the topics sometimes varies from G99.

**Measurement of performance**

Measurement of performance has dominated performance research (Figure 1). In G99, the distribution of papers dealing with measurement of various performance variables is as shown in Figure 2. Most papers pertained to timing and dynamics as well as to intonation and vibrato; the large number of papers in the latter category mainly reflected the numerous papers on vibrato from the Iowa research during the 1930s. Timing, in particular, has been the primary subject for investigation, which is to be expected as timing is used in performance on all instruments as well as singing, whereas variation of dynamics and intonation is not possible or is limited with certain instruments (e.g. piano, organ, harpsichord). The term ‘timing’ usually refers to how the duration of single notes or other entities ‘deviates’ from a norm, such as a ‘mechanical’ or ‘dead-pan’ performance with absolutely constant tempo and

![Figure 2](image-url)
strict adherence to the ratios between note values in the score. Such deviations were reported from the very beginning of performance research, and numerous examples in G99 demonstrated different types of timing, at different levels (e.g. at note, beat, bar and phrase level) and in different kinds of music.

TIMING AND DYNAMICS
Measurement of performance continues to be a dominant theme in recent research. Extensive measurements of timing and dynamics in piano or keyboard performance have been reported by Repp. An earlier study of 24 famous pianists’ performance of *Träumerei* (Repp, 1992a; see G99) was followed by studies of graduate student pianists’ performance of the same piece (Repp, 1995a). In terms of group average timing pattern and individual consistency, students turned out to be comparable to the famous pianists but much more homogeneous among themselves than the experts who demonstrated more interpretive originality; similar results were also obtained in a comparison of students’ and expert pianists’ timing in Debussy’s *La fille aux cheveux du lin* (Repp, 1997a). One important reason for the relative uniformity among young performers is probably the highly competitive nature of today’s music business (see also discussion in Repp, 2000). Students’ use of dynamics in *Träumerei* was similar and consistent over repeated performances (Repp, 1996a).

The performance of Chopin’s well-known *Etude in E Major* (op. 10, no. 3, bars 1–5) was likewise studied with regard to timing (Repp, 1998a) and dynamics (Repp, 1999a), using no less than 115 commercially recorded performances spanning 68 years, from 1927 to 1995. Factor analysis (principal components) revealed four major timing strategies related to ritards at the ends of melodic gestures, acceleration within some gestures, lengthening of the initial downbeat, and other ritards. However, only a few pianists’ performance conformed closely to any of these strategies, most performances represented some kind of weighted combination of different strategies plus idiosyncratic variation. Likewise, most performances showed dynamic patterns representing some kind of mixture of different dynamic profiles found by factor analysis. Timing and dynamics seemed largely independent of each other. The overall aesthetic quality of the performances was assessed by four judges specially selected for the research (Repp, 1999b). It turned out that timing and dynamics, usually considered the most expressive performance variables, accounted for only a small part (10–18%) of the variance in the judges’ ratings. It was speculated that other variables – for instance, the elusive variable called ‘touch’ in piano performance – might have greater import than timing and dynamics; nor were the conditions for the assessment optimal. All three studies were summarized in Repp (1999c). Pianists’ attempts to imitate various, less typical model performances of the Chopin piece were influenced by their cognitive schema of the typical, spontaneous expressive pattern for this piece (Repp, 2000).
Repp (1996d) analysed graduate student pianists’ pitch errors in performances of four pieces. Errors were classified exhaustively as substitutions, omissions and intrusions. They were concentrated in technically difficult parts of the piece and occurred almost always in non-melody voices, often inside chords. Most errors were contextually appropriate and were thus not noticed by listeners.

Further studies included a study of pedal timing in *Träumerei* (Repp, 1996b) and timing of arpeggiated chords in one of Grieg’s lyric pieces for piano, *Erotik* (Repp, 1997c). Pedal timing by a concert pianist was also studied by Palmer (1996a, 1996b); the pianist consistently used pedal releases before note onsets to prevent dissonances across successive events. Both the use of the pedal and arpeggio performance were little studied in earlier research. Generally, all these papers provide much useful information and discussion regarding problems of measurement, definitions of concepts and interpretation of findings.

In a comparison between individual performances and an ‘average’ performance (Repp, 1997b) – realized as an average of 11 student performances of *Träumerei* or as the average of 30 synthesized performances of the Chopin excerpt varying in timing and tempo – listeners judged an average performance as best or among the best with regard to quality but at the same time lacking in individuality; there was a negative correlation between quality rankings and ratings of individuality. An interesting paradox was that, although the average of the experts’ performances was ranked highest in quality, most of the single expert performances were ranked lower than performances by students. It was suggested that one possible reason, among many others, could be that expert pianists deliberately distorted timing in order to provide a new, unexpected shape to this so frequently played Chopin piece, even at the expense of perceived beauty according to conventional standards.

Inspired by research in Uppsala, Humair (1999) conducted a comprehensive study to investigate systematic variations (SYVAR) in different types of dance music (bourrée, habanera, march, mazurka, minuet, ragtime, samba, sarabande, swing, tambourin, tango, waltz, Viennese waltz) by analysing professional pianists’ performances of such dances on a Yamaha Disclavier. The existence of different SYVAR patterns in timing on note level and beat level was amply confirmed. His dissertation also included a thorough historical survey of various authors’ views on *notes inégales* and tempo rubato.

The relationship between timing (tempo) and dynamics in performance has been studied and discussed in several papers with special reference to Todd’s (1992) model concerning a coupling between these two variables (‘the faster, the louder’ and vice versa). Results in studies by Clarke and Windsor (2000), Palmer (1996a), Repp (1999a), and Windsor and Clarke (1997) indicated that this version of the model was unsatisfactory in some respects, and that a simple positive correlation between tempo and dynamics was not...
adequate as a general description of their relationship. Sometimes timing and dynamics seem independent of each other; sometimes they may have similar functions, and at still other times, they may interact in different ways.

The question of whether microstructure timing is dependent or not on tempo is still open (see G99: 541). Repp et al. (2002) studied this in the performance of simple musical rhythms and found that the timing of two-note rhythms changed little with tempo. With three-note rhythms, the short interval was mainly unaffected, but the two longer intervals showed assimilation as tempo increased; this result thus contradicted the relational invariance hypothesis that relations between note durations remain the same at different tempos. Another finding that contradicts this hypothesis is that the 'swing ratio' in jazz performance decreases with increased tempo (Friberg and Sundström, 2002). This is discussed in more detail later in this article.

Palmer (1996b, see also 1996a) studied so-called melody lead in piano performance: melody notes in chords come somewhat earlier (20–50 ms) than other notes in the chord; this effect was more pronounced for expert performers than for student performers. Sometimes melody lag was used as well. It has been questioned (Goebl, 2001; Repp, 1996c) as to whether melody lead is a genuine timing device used in pianists’ key-presses, or whether it is an artifact due to the melody note being played louder (with higher velocity than the other notes in the chord) thus generating earlier hammer-string contact for this voice, although all notes in the chord were played simultaneously (at the keys). However, these alternatives need not necessarily exclude each other but could occur in different ways in different contexts.

Most performance studies deal with piano or keyboard performance. Johnson (1996a, 1996b) analysed timing in four performances of the first movement of Mozart’s *Concerto No. 2 in E-flat Major for Horn and Orchestra*. Two of these performances were judged by experts as exceptional and demonstrated much more use of rubato than two other, inferior performances. Johnson (1997) also asked five professional musicians to assess perceived musicianship for each note of these four performances by placing a rating from –5 to +5 on the corresponding note in the score; ratings tended to increase in conjunction with increased use of rubato. Positive effects of instruction concerning appropriate use of rubato were demonstrated (Johnson, 1998, 2000).

Johnson (1999) also measured timing in 15 performances of Bach’s *Suite No. 3 for Violoncello Solo, Bourrée No. 1* and found typical acceleration–deceleration patterns within phrases. There were further variations in timing due to instrument specific challenges (e.g. performance of triple stops, wide intervals). In another study (Johnson, 2000–1), musical experts, using a MIDI keyboard in combination with a special computer program, performed the thematic section of Beethoven’s *Symphony No. 5* in three versions: mechanical, interpreted and exaggerated. Analysis of these performances showed that the three versions differed considerably more in timing than in dynamics.
As pointed out earlier, the bulk of performance measurements pertain to western classical music; so far, there is little on other musical genres. Prögler (1995) discussed the concept of swing in jazz music, as described by jazz musicians and writers on jazz, and measured the timing of drummers’ ride taps in relation to a beat provided by a metronome or a recorded bassline. While one drummer usually played the ride tap ahead of the bassline, another drummer played it after the bassline, and sometimes the drummers alternated between playing ahead or after the beat.

A recent issue of the journal *Music Perception* (vol. 19[3]) was devoted to jazz performance. Performances by jazz giants such as Louis Armstrong (Collier and Collier, 2002), Chet Baker, Art Farmer, Miles Davis and John Coltrane (Ashley, 2002), Tony Williams and Jack DeJohnette (Friberg and Sundström, 2002), Thelonius Monk and Ahmad Jamal (Iyer, 2002), and various jazz pianists (Busse, 2002) were analyzed regarding tempo and timing. Some results are briefly described here. Tempo (beat timing) was very regular; the ability to keep time is one of the key requirements for good jazz musicians, especially those in the rhythm section. Unlike the small and barely perceptible melody lead in classical music, melody notes in jazz performance are rather delayed in relation to the accompaniment – the delay is usually clearly perceptible and can at times amount to as much as one or more beats, reflecting the freedom in jazz performance – but accelerate during the course of the motif or phrase to align with the accompaniment at cadential locations (Ashley, 2002). The so-called ‘swing ratio’ between two successive notes of nominally the same duration (if notated at all) is usually taken to be a 2:1 ratio (‘triple feel’; see G99: 537). Data from these references (see also Prögler, 1995) showed that this ratio may in fact vary from about 4:1 to 1:1, depending on the performer and the context. For drummers, Friberg and Sundström (2002) found an approximately linear decrease of the ratio with increased tempo, from about 3:1 at 100 bpm to about 1:1 at 300 bpm. The ratio for the soloist was always lower, typically between 2:1 and 1:1, than for the drummer. This difference was possible because the soloist was delayed at downbeats but synchronized at off-beats. Iyer (2002) even described a performance case with ‘inverted’ swing ratio. A special problem in analyses of jazz performance is that, when there is no score available to act as a ‘norm’, results may differ depending on what transcription of the performance is used (see, e.g., Ashley, 2002). This problem also appears in other contexts, for example, in comparison between transcriptions and measurements of performances of Cuban dances (Alén, 1995).

Finally, just a brief mention of other papers on measurement of various performance parameters: performance of legato, staccato and repeated notes (Bresin and Battel, 2000); a neural-network model for control of legato articulation in playing scales and arpeggios (Jacobs and Bullock, 1998); timing in Estonian folk songs (Ross and Lehiste, 1998); performance of ornaments (Timmers et al., 2002; see also Palmer, 1996a); and performance timing in
different contexts: the playing of a Brahms melody with bar-lines removed, with bar-lines, together with counter-melody, together with block chords, and in the complete setting (Timmers et al., 2000).

INTONATION, VIBRATO

Most early studies on intonation focused on intonation in relation to different tuning systems (just intonation, Pythagorean tuning, equal temperament) and on intonation dependent on acoustical and expressive purposes (G99: 545–7). Basic results concerning rate and extent of vibrato were obtained in the Iowa research (G99: 529–31, 546–7).

There seem to be few new studies on intonation and vibrato. Prame (1997) analysed 10 singers’ performance of Schubert’s Ave Maria. Mean vibrato extent for individual tones was between ±34 and ±123 cent, and artists’ means varied between ±57 and ±86 cent; the longer the tone, the smaller the vibrato extent. Violinists performing the same piece used much smaller vibrato, half or less of the extent of the singers’ vibrato. Singers’ intonation showed remarkable departures from equally tempered tuning, the greatest tone mean departures amounted to +42 and −44 cent. Within a singer the greatest difference between sharpest and flattest tone mean was 69 cent. An earlier study (Prame, 1994; G99: 546) analysed singers’ vibrato rate.

A review of research on vocal vibrato was made by Sundberg (1995); see also Sundberg (1999) for a general review of research on singing.

Models related to measurements

Measurements of performances result in large amounts of data, even for short pieces. There is an urgent need to find some principles that may summarize and suggest explanations of the phenomena discovered (e.g. the ‘deviations’ in timing). Seashore (1938: 29) proposed that musical art generally relies on ‘artistic deviation from the fixed and regular: from rigid pitch, uniform intensity, fixed rhythm, pure tone and perfect harmony’. Similar reasoning underlies later formulations in terms of ‘systematic variations’ or ‘expressive deviations’. However, these principles are too general to have good explanatory value.

More precise proposals have appeared in recent decades. Some models were discussed in G99 (pp. 550–6): models based on measurements (Clarke, 1988; Todd, 1989, 1992, 1995) or on intuitions (e.g. Clynes, 1987; Friberg, 1995; Sundberg et al., 1991). These are still relevant (partly revised), and more proposals have emerged as described later.

A series of papers by Repp dealt with listeners’ ability to detect increments and decrements in timing of notes in different positions. For instance, increment of a note interonset interval (IOI) in mechanically timed music proved to be more difficult to detect in positions where the note is typically...
lengthened in musical performance – a kind of perceptual bias first demonstrated in Repp (1992b) and further investigated in Repp (1998b, 1998c). Listeners’ timing expectations seemed about five times smaller than typical expressive timing variations and are thus essentially subliminal. A related phenomenon is that, when instructed to play in a strictly mechanical way, musicians still unintentionally exhibit timing variations of the same type as in expressive performance but on a reduced scale (e.g. Bengtsson and Gabrielsson, 1983; Palmer, 1989; Penel and Drake, 1998; Repp, 1999f).

The origin of these phenomena may be sought in different places: in listeners’ learned expectations, in various psycho-acoustical/perceptual factors (e.g. interactions of pitch, intensity, duration), in motor constraints and in the demands of the musical structure itself, such as temporal implications of melodic-rhythmic groups at different levels and positions in the music, or in varying combinations of all these factors, depending on the musical context. These matters have been discussed in several papers (Penel and Drake, 1998, 1999; Repp, 1995b, 1998d, 1999d, 1999e, 1999f). Penel and Drake (1999) discussed three attempts at explanation of expressive timing in music performance: (a) use of timing to highlight and communicate the musical structure to the listener, so far the most common hypothesis; (b) a ‘perceptual’ hypothesis meaning that some timing variations are due to compensations for perceptual biases in time perception (e.g. psycho-acoustic effects of pitch and intensity on time perception and effects related to rhythmic and melodic grouping); and (c) a ‘motor’ hypothesis meaning that timing variations may be due to biomechanical and instrument-related constraints (e.g. regarding fingering in piano performance; for an example in drumming performance, see Alén, 1995) and on the existence of certain preferred, ‘natural’ motor patterns. After discussing the pros and cons of these alternatives it was concluded that all of them may contribute, in different combinations, to timing variations. A paradigm to investigate their relative contributions was proposed. Ask the participants to: first, adjust the durations (IOI) in a sequence to achieve perfect regularity (only perceptual factors involved); second, perform a perfectly regular sequence (involving both perceptual and motor factors); and finally, perform the sequence in a musical way (adding ‘expression’).

A further factor that may affect timing in addition to other performance variables is performers’ intention to express certain emotions. After all, expression in music is predominantly associated with emotional expression (Gabrielsson, in press a; Gabrielsson and Juslin, 2003). While this factor has been neglected in earlier performance research, a large number of papers on emotional expression in performance have appeared recently (Bresin and Friberg, 2000; De Poli et al., 1998; Gabrielsson, 1994, 1995, 1999b; Gabrielsson and Juslin, 1996; Gabrielsson and Lindström, 1995; Juslin, 1997a, 1997b, 2000, 2001; Juslin and Laukka, 2000; Juslin and Madison, 1999; Laukka and Gabrielsson, 2000; Madison, 2000; Mergl et al., 1998;
Rapoport, 1996; Senju and Ohgushi, 1987; Siegwart and Scherer, 1995; Sundberg, 2000; Sundberg et al., 1995), most of them reviewed in Gabrielsson and Juslin (2003). In most of these studies performers were instructed to play/sing the same musical piece, usually monophonic, so as to express different emotions, such as happiness, sadness, anger and fear; sometimes an ‘expressionless’ performance was asked for as well. A variety of different instruments – drums, electric guitar, flute, keyboard, piano, violin and singing voice – have been used. Analysis of the recorded performances revealed that practically every performance variable – tempo, timing, dynamics, intonation, articulation, vibrato, tone envelopes, timbre, etc. – was affected in ways specific to each emotion; for summaries, see Gabrielsson and Juslin (2003), Juslin (2001) and Juslin et al. (2001–2).

This and other research has recently generated a new, comprehensive model of expression in music performance, the GERM model (Juslin et al., 2001–2). This constitutes an attempt to integrate four possible sources of microstructure in music performance into a common model: (a) generative rules (referring to musical structure), (b) emotional expression, (c) random variations (reflecting internal time-keeper variance and motor delay variance), and (d) movement principles (assuming that certain performance features are shaped according to principles in biological motion). The model is designed as a computational model to permit simulation of different aspects of performance expression and empirical test of the validity of the model as a whole or of its different components. Evidence and ideas from many different research areas – music theory, speech perception and production, emotion perception and emotion theory, brain mechanisms, time-keeper theory, biological motion and others – were invoked as support for the model, and a first experimental evaluation of the model has been reported. Further description and discussion of this model appears in the paper by Juslin (2003, this issue) on five facets of musical expression.

The comprehensive KTH rule system (e.g. Friberg, 1991, 1995; Sundberg et al., 1991) discussed in G99 (pp. 552–4) has been supplemented with new rules regarding punctuation (marking of small melodic units by lengthening of the last note followed by a micropause; Friberg et al., 1998): rules regarding tone duration and sound level in combination with other rules in order to express different emotions (Bresin and Friberg, 2000; cf. also Juslin et al., 2001–2). The rules are implemented in the Director Musices program (Friberg et al., 2000). Furthermore, Bresin (1998, 2000) described a model based on artificial neural networks (ANN) to learn some of the KTH performance rules and the style of a professional pianist; listening tests indicated that the results were musically acceptable.

Another attempt at establishing rules for music performance is represented by a comprehensive project conducted by Widmer (2001, 2002) in Vienna. The project aimed to develop computational methods to study expressive music performance and to inductively build formal models of
expressive performance using intelligent data analysis methods from Artificial Intelligence (AI) research and machine learning in order to detect patterns and regularities in many examples of high-class pianists' performances on the Bösendorfer SE 290 grand piano equipped with a special mechanism for recording key and pedal movements. The so-called PLCG Rule Discovery Algorithm (Partition + Learning + Cluster + Generalize) is used to search for classification rules that, in a set of positive and negative instances of a phenomenon (e.g. notes played staccato), describe as many of the positive instances as possible while covering as few of the negative instances as possible (e.g. non staccato). Applying this to data on timing, dynamics and articulation in recordings of 13 complete Mozart piano sonatas and some Chopin pieces resulted in a tentative set of rules for categories in timing: lengthening–shortening of IOIs, dynamics (louder–softer) and articulation (staccato–legato) with varying degrees of coverage and precision. In addition to such note-level rules, work has started to build models at higher levels, for example, at phrase level (Widmer and Tobudic, in press). It is emphasized that, unlike music performance researchers who may have various preferences for which kind of explanatory principles should be tried, such an inductive machine learning approach has no theoretical ‘bias’.

Another proposal for measurement and representation of performance data at different levels of the musical structure is based upon algorithms of the Theory of Oscillating Systems (Langner, 2000, 2002; Langner et al., 1998). Basically, a large set of oscillators, each with a fixed frequency and phase and arranged in logarithmic steps from 8 Hz down to 0.008 Hz (corresponding to MM 480 to 0.48) is used to detect periodicities in the loudness curve – expressed in sones according to Zwicker’s model of loudness (Zwicker and Fastl, 1990) – of a music performance. Each oscillator contains an activation window which opens and closes in accordance with the respective frequency and phase. If a musical event occurs while the window is open, it will activate the oscillator. The result is displayed in an oscillogram with time along the horizontal axis, and frequency or corresponding metronomic value along the vertical axis. The periodicities appear as shadings, at the detected time-frequency positions, within the oscillogram, which allows study of the timing at several different temporal levels. Many instructive examples were given in the papers referred to earlier. A partly analogous procedure is used to display variations in dynamics in so-called dynagrams. The multiscale model of rhythmic grouping proposed by Todd (1994) has some similar features in assuming that the auditory system operates as a number of energy-integrating low pass filters with different time constants. Langner and Goebl (in press) also recently demonstrated a real-time visualization of a performance in tempo–loudness space. As the music goes on, a dot moves through this space leaving behind it a trajectory (a ‘worm’) that elucidates the interaction between tempo shaping and dynamics.

The model proposed by Mazzola and Beran (1998) may be the most
mathematically inclined model; the authors claimed that the transformation from score to performance should be described in ‘effective mathematical terms’ (p. 40). Starting from a musicological analysis of the underlying score regarding rhythmical, melodic and harmonic structure and calculation of weight functions for these parameters, an agogical operator generates the timing microstructure as function of the weights, all conducted by the so-called RUBATO analysis and performance workstation. Regression analyses were used to study how well the model fits the data. As an example, the procedure was applied to Schumann’s *Träumer*, with Repp’s (1992a) measurements of performances of this piece available as reference. The mathematics of this model may be beyond what most music psychologists (including me) can follow. Perhaps the most interesting point is the authors’ emphasis on score analysis as a basic means for expressive semantics, in other words, for ‘rational composition of performance’. Another, less developed, approach (Tangian, 1999) similarly took structural analysis as its starting point and proposed a partition of a piece into segments, at different levels, each of which is performed with a certain ‘tension curve’ through timing and dynamics.

Clarke (1995), by way of introduction, discussed three different theoretical perspectives on expression: (a) generative theory, i.e. that expressive performance includes systematic patterns of deviation from the ‘neutral’ information in the musical score; some examples would be the models proposed by Clarke (1988), Mazzola and Beran (1998), or Tangian (1999); (b) expression as integrated energy flux as proposed in a multiscale model of rhythmic grouping (Todd, 1994) mentioned earlier; and (c) narration/drama in performance expression, emphasizing the necessity of narration and emotional expression as underlined by Shaffer (1992, 1995; see also G99: 503). Clarke (1995) then proposed a semiotic approach based on Peirce’s (1931–58) explications of three kinds of signs: index, icon and symbol. In performance, index and icon seem more important than symbolic signification which, on the other hand, is crucial for signification in musical structure. An example of iconic signification are ritards at phrase endings to signify the phrase boundary’s structural importance; an example of indexical signification may be how the sound quality of a performance reflects physical effort in producing the sound, for instance, in coloratura performance. These ideas were examined in detailed analyses of different performances of Chopin and Beethoven pieces. Within the constraints imposed by the musical structure, the performer can manipulate acoustical and temporal variables to lead, or mislead, the listener in an evolving interpretation of the structure of the piece and in extra-musical characterizations; for instance, extra-musical narrative. The various expressive features (instances of timing, dynamics) may serve as clues to different interpretations in ways that may not be understandable if they are considered in a too myopic way: ‘surface features point well beyond their immediate context to fundamental differences of conception’ (p. 50).
There are many other models also connected with performance of certain music or certain musical components; for instance, models for bebop improvisation (Johnson-Laird, 1991, 2002; Toiviainen, 1995; see the Improvisation section of this article) and for performance of musical ornaments (Timmers et al., 2002). Models in connection with other topics in performance appear sporadically in the rest of this article.

Some comments on measurements and models

A perennial problem in music performance research is the amount of data obtained in measurements. New approaches present improved facilities for collecting and analysing large amounts of data (e.g. Langner, 2002; Widmer, 2001, 2002), and different forms of multivariate statistics (e.g. factor analysis, multiple regression) are now routinely used. The focus is mostly on tempo and timing in performance, but other aspects – dynamics, intonation, articulation, vibrato, tone envelopes, timbre, pedalling, as well as possible interactions between variables – have been given increased attention. Piano/keyboard performance remains the most studied topic, not least due to readily available technical facilities (e.g. MIDI) for the recording of performance as well as for synthesizing stimuli for listeners’ judgements. As different instruments offer different possibilities in performance, there is a need for an extended repertoire of instruments in performance research. And, of course, performance research has to broaden its scope beyond western art music that is currently by far the most predominant area for performance research.

A number of models for descriptive and explanatory purposes have been suggested as briefly described earlier. They vary considerably in general background, aim and direction, degree of formalization, scope and possibilities for empirical testing. Some of them depart from theoretical constructs in physics and physiology (Langner, 2002; Todd, 1994), others strive to work in a purely inductive way (Widmer, 2001), or formulate rules based on experienced musicians’ intuitions (Friberg, 1995; Sundberg et al., 1991). In addition, others relate, in various ways, to processes in perception and motor functioning (Penel and Drake, 1999), biological motion, emotion, brain functions and others (Juslin et al., 2001–2). In all of these proposals there is, of course, also varying reference to musical structure as manifest in scores, most emphasized in the approach by Mazzola and Beran (1998). The semiotic approach suggested by Clarke (1995) is perhaps the most ‘humanistic’ alternative and relies heavily on human interpretation of various signs in the structure and performance of music.

It remains to be seen how the various models stand up to empirical testing. Of course, none of them can represent a model of music performance in a more general sense. An all-inclusive model would have to take a lot of other aspects into consideration, such as acoustical conditions, stylistic conventions, performance practices, performers’ skills and personalities, social
processes in connection with performance, and so on. This is a goal that is a long way off and that can be reached only through the accumulated efforts of many researchers.

To conclude this section, I quote from G99:

> It is my conviction that measurements of performance should, as much as possible, be conducted and considered in relation to the composer’s and/or the performer’s intentions and the listener’s experience ... After all, music is a means for communication and expression, and the characteristics of different performances may be easier to understand given this self-evident frame of reference. (p. 550)

**Planning of performance**

While measurement of performance has been the predominant activity in empirical performance research since its very beginning around 1900, most other areas (see Table 1) were not investigated until much later. Studies on planning of performance – that is, how to form mental representations of the music, devise performance plans and strategies for efficient practice – did not start until after the emergence of cognitive psychology during the latter half of the 20th century; there was little place for such questions as long as behaviourism dominated psychology. Exceptions, however, were the pioneering studies on different methods of memorizing music (massed or distributed practice, whole or part learning, overlearning, etc.) conducted by Rubin Rabson in the 1940s (see G99: 507).

Mental representation of a piece of music may refer to many aspects of the music, varying in different situations and with different performers. Most empirical studies so far have focused on performers’ representation of the structure of the music, especially on the relation between different levels within an hierarchical structure. Successively, other aspects have gained increased attention, such as the performer’s conception of feelings and emotions, meanings, narratives and movements (imagined and real) associated with the music (G99: 502–5).

In a continuation of an earlier study (Palmer and Van de Sande, 1993) on units in representation of homophonic and polyphonic music, Palmer and Van de Sande (1995) investigated the range of planning in music performance using data concerning the frequency of errors in performances (e.g. substitutions, deletions, additions of notes) and timing as indicators. It was hypothesized that the range of planning would be affected by the serial distance between notes (nearby notes should contribute to planning more than notes farther away) and by the musical structure (elements that belong to the same phrase are more likely to be planned together). Advanced pianists performed two- and three-part inventions by Bach, in which the entrance of the second voice was either ‘early’ (7–8 notes after the first voice entrance) or ‘late’ (11–12 notes after) in three different phrasing alternatives. Results
suggested, as expected, that the range of planning was affected by both serial distance and phrase structure. Moreover, there was an interaction between these factors; for instance, performances of notations that lacked phrase boundaries led to more errors when the second voice entered ‘late’, whereas performances of the same music with notated phrase boundaries showed more errors for ‘early’ entrance of the second voice.

Analysis of errors in performance was also used in experiments with skilled pianists and novice child pianists to test a detailed model for incremental planning in sequence production (Palmer and Pfordresher, in press). The model builds on the assumptions that events (notes) are encoded in terms of their serial order and timing relative to other events in a planning increment, and that planning is facilitated by the metrical similarity and serial proximity of the events, further by developmental changes in short-term memory. The model’s predictions of larger planning increments with decreasing production rate and with increasing age and experience of the performers were confirmed.

Another aspect concerns what musicians focus on in planning performance. Sullivan and Cantwell (1999) used path analyses to model how different factors affected planning focus (divided into seven levels, where the highest level represented focus on the meaning and interpretation of the music) in university music students confronted with a traditional score and a non-traditional score (graphic notation) presented one line at a time on a computer screen. After a completed reading of the score, the participants verbalized how they would go about learning the score to a level of performance competence. Components in the model included: a deep vs surface approach to learning; reading times and reaction times indicating depth of cognitive engagement; and low-, mid- and high-level planning strategies. Use of low-level strategies meant, for instance, rote learning, trial-and-error and sight-reading; mid-level strategies included speed alteration, chunking and linking of elements; and high-level strategies included interpretation, patterning, prioritizing and monitoring. With regard to the traditional score, the resulting model showed a strong relation between high-level strategies and (high-level) planning focus; high-level strategies were in turn positively related to a deep approach to learning and to time used in reading the score. The result for the non-traditional score was partly similar and also included a component representing prior familiarity with graphic notation.

PRACTICE

Studies of musicians’ rehearsal techniques did not start until the 1980s (e.g. Gruson, 1988; Miklaszewski, 1989; see G99: 508) and are now increasingly researched and facilitated thanks to the possibilities offered by audio or video recordings (G99: 508–9).

Some studies rely on verbal reports by performers obtained in interviews or from questionnaires (e.g. Hallam, 1995; McPherson and Cormick, 1999).
However, most of the studies mentioned in the following paragraphs used recordings (audio or video) of practice behaviour in combination with some kind of verbal report. The verbal reports come from performers instructed to ‘think aloud’ during the practice and/or from interviews afterwards. For instance, Nielsen (2001) asked performers to think aloud as if answering questions such as ‘What am I thinking?’ and ‘What am I focusing on?’; she also interviewed the performers afterwards, asking them to look at the video recording to stimulate recall of their thinking, a procedure also used earlier by Miklaszewski (1995). The combination of recording and verbal reports is generally preferable if conditions allow. Of course, thinking aloud may interfere with performance itself and performers may find it difficult (Williamon et al., 2002). Observation of practice behaviour can provide important information that does not appear in the performer’s verbal reports. For instance, although observation revealed that the pianist practised dynamics right from the start, this was not mentioned in her self-reports, probably because it was unproblematic (Chafin and Imreh, 2001).

The music used typically belonged to western art music. The performers were pianists (e.g. Chafin and Imreh, 2001; Miklaszewski, 1995; Williamon and Valentine, 2002), organists (Nielsen, 1997, 2001), string players (Hallam, 1997a, 1997b), orchestral musicians playing different instruments (Hallam, 1995), or singers (Ginsborg, 2002).

All the researchers found marked individual differences in practice behaviour. Following her interviews with 22 orchestra musicians, Hallam (1995) distinguished between analytic holists, intuitive serialists and versatile learners. The first two categories were somewhat unexpected since an analytic approach is usually thought to be combined with a serialist strategy (that is, working through the music sequentially). The largest category, versatile learners, adopted holist and serialist strategies interchangeably. The performers in Miklazewski’s (1995) study may as well be described as representing holists, serialists and versatile learners.

At some point during practice, right from the start or later, performers usually segment the piece into parts corresponding to different levels of the musical structure and practise them separately. However, some parts may be selected mainly because they present technical difficulties; such parts may be as short as a single bar or a few bars. Successively the length of segments to be practised becomes longer and longer to be finally integrated in a complete performance. Naturally, work on technical problems occurs early in practice and then decreases in favour of matters of interpretation. These findings recur in practically all studies. For instance, Williamon and Valentine (2002) found that pianists, especially those at higher levels of skill, increasingly started and stopped their practice on important structural bars, while starts and stops on difficult bars successively decreased. In the study by Chafin and Imreh (2001; Chafin et al., 2002), practice of the Presto movement in Bach’s Italian Concerto included 57 sessions over 10 months, totalling about
30 hours of practice. The pianist made her comments as she practised and after the performance also reported her decisions on three ‘basic’ dimensions (fingering, technical difficulties, familiar patterns of notes), four interpretative dimensions (phrasing, dynamics, tempo, pedalling) and three performance dimensions representing features attended to during performance (basic, interpretative, expressive). Her practice was divided into three periods. Practice was affected by the basic dimensions in the first two periods and by the interpretative dimensions during the last two periods, while performance dimensions affected practice throughout. Williamon et al. (2002) found that pianists successively increased the length of practice segments but that they also increasingly interspersed short and long segments; this was especially evident for performers at higher levels of ability. This suggested that skilled performers shift their focus of attention between different levels of the musical structure in accordance with a suggestion made by Clarke (1988).

How practice is conducted depends on the performer’s level of musical skill. Hallam (1997a, 1997b) compared professional musicians with young string players (aged 6–18 years) concerning the regularity and extent of practice; structure and routine of practice; learning of unfamiliar music; identification of difficult sections; development of interpretation; memorization of music; and preparation for performance. The professionals demonstrated extensive meta-cognitive abilities: self-awareness of their own strengths and weaknesses; knowledge of the nature and requirements of different tasks and of strategies to complete them satisfactorily through planning, monitoring and evaluation of practice and performance. Findings for the most advanced students were similar to those of the professionals, although less definite. However, practically none of the novices showed or reported any activities that could be related to interpretation; ‘the aim of the practice was to play the music correctly’ (Hallam, 1997a: 99), especially with regard to pitch at the expense of rhythm. Recordings of their practising a short piece showed various strategies: just playing through the whole piece several times, playing through but stopping meanwhile to practise larger sections, playing ‘one line at a time’. Errors were left uncorrected or led to the repetition of a single note or a small section (cf. Gruson, 1988).

While memorizing of music both for professionals and students occurred through automated processes such as visualizing the music (e.g. ‘seeing’ actual notes), relying on aural memory of the music, or through kinaesthetic/motor memory (‘fingers sort of know where to go’), professionals also used analysis of the music’s structure to assist in memorization. Likewise Williamon and Valentine (2002) concluded from their findings that identification and use of musical structure is a salient characteristic of skill both in guiding practice and in serving retrieval during performance.

However, in a study on female singers’ practice in learning an unknown song (Ginsborg, 2002), experienced professional singers were found to be neither faster nor more accurate memorizers than students and amateur
singers. Ginsborg instead formed a subgroup of fast and accurate singers taken from all three categories. In comparison with another subgroup of slow, inaccurate memorizers, fast and accurate memorizers were more likely to produce music and words together in their first initial practice sessions; they began memorizing earlier and counted beats aloud during the learning process; used more different modes in learning – such as speaking words, singing words, vocalizing, playing the melody, accompanying; they made as many errors, mainly pitch and rhythm errors, as slow, inaccurate singers but both made and corrected them earlier. There were thus clear differences between the two subgroups, although they did not reflect a division between professionals and students/amateurs.

Hallam’s (1997a, 1997b) findings indicated that professionals learn to learn, that is, to develop appropriate strategies for practice and performance. McPherson and McCormick (1999) found that harder working students reported higher levels of cognitive strategy use, such as rehearsing music in their minds, critically evaluating their efforts, and organizing their practice to achieve efficient learning; students who were more cognitively engaged while practising seemed to not only practise more but also to be more efficient with their learning. Which strategies are most appropriate of course depends very much on the type of music, on the specific instrument, on the performer’s level of skill, etc. For example, in her study of two organ students, Nielsen (1999b, 2001) discussed selection strategies (selection of problem areas), strategies for organizing the material (e.g. dividing the piece into different ‘working areas’, writing markings in the score) and the work (e.g. working with and repeating material in segments of different length, in different tempos, in uni- and bilateral playing, with different accent structures), and strategies for integration.

Generally, then, cognitive engagement during practice seems an important key for improving practice and refining musical skills. This is evident in other studies as well. In a transfer learning task designed to study the relative influence of conceptual and motor factors (Palmer and Meyer, 2000), child pianists’ transfer to the new task was affected by both conceptual (if the pitch sequence was the same or different) and motor (if fingering was the same or different) factors. For skilled pianists, however, transfer depended only on conceptual similarity or dissimilarity; motor factors (fingering) did not matter. These findings suggest that, with increasing skill, mental representations for performance become successively more dissociated from the movements involved. (In fact, highly skilled pianists may do their practice away from the piano; see G99: 506).

Drake and Palmer (2000) studied learning of a novel piece by five groups of pianists differing in age, skill and sight-reading ability. Performance tempo and accuracy (= less errors) increased with skill level and practice in early learning trials. Making pauses (breaks) were the most frequent errors, especially in the less skilled groups, followed by pitch and correction errors, and
timing errors, in that order. The less skilled groups made a higher proportion of timing errors than pitch errors, focusing on ‘what’ to play rather than ‘when’ (cf. Hallam, 1997a), whereas the more advanced pianists had a higher proportion of pitch errors relative to timing errors (focusing on ‘when’). Temporal continuity – that is, avoiding pauses, corrections, etc. – and anticipatory errors, which indicate future-oriented planning, increased with both skill level and practice. Range of planning, measured as distance between interacting elements in pitch errors (cf. Palmer and Van de Sande, 1995), increased with skill level. It was suggested that the increased mastering of temporal constraints and increased planning abilities reflect the same underlying cognitive processes, relating to factors such as segmentation and size of segments, observation of hierarchical structure and ability for simultaneous planning and execution. In a longitudinal study (3 years) of high-school instrumentalists, McPherson (1997, 1999) found increased skill in sight reading, playing by ear, playing from memory and improvising music, also in the development of cognitive strategies for these activities; however, individual differences were large.

DELIBERATE PRACTICE
A series of papers (Ericsson, 1997; Ericsson et al., 1993; Krampe, 1997; Lehmann, 1997a, 1997b; Lehmann and Ericsson, 1993, 1996, 1997, 1998a, 1998b) concern the use of so-called deliberate practice to achieve high-level music performance. It is claimed that the main factor behind expert performance is the amount of extended (at least 10 years) and optimally distributed deliberate practice. Deliberate practice means carefully structured activities in order to improve performance and presupposes high motivation and extended effort, full attention during practice (which limits the length of practice sessions and necessitates time for recovery); explicit instructions and individualized supervision by a teacher; knowledge of results; favourable environmental conditions; and parental or other support. This theme was illustrated and elaborated with regard to sight reading and accompanying performance (Lehmann and Ericsson, 1993, 1996), a piano soloist’s practice of pieces before concert performance (Lehmann and Ericsson, 1998a), and with regard to implications for music education of amateur musicians and music students (Lehmann and Ericsson, 1997), also in a survey of historical development of expert performance (Lehmann and Ericsson, 1998b). Lehmann (1997a, 1997b) further discussed three necessary mental representations: (a) mental representation of the desired performance goal; (b) mental representation of the current performance; and (c) representation of the music with regard to its production aspects – and how these pertain to accompanists and soloists, respectively.

A large-scale English study (Sloboda et al., 1996), including five groups of young people (8 to 18 years old) differing in attained musical achievement, showed that there was indeed a strong relationship between musical
achievement and the amount of formal effortful practice (and a weaker relationship between achievement and informal training). Probably, nobody doubts the importance of much extended practice. However, Williamson and Valentine (2000) provided some evidence against a simple monotonic relation between amount of practice and performance quality. They divided 22 pianists into four levels of skill and asked them to practise pieces of Bach to be performed in public. Their practice was recorded on tape and the amount of practice could be calculated. Sheer amount of practice was not significantly related to performance quality of the final performances as judged by three experienced judges regarding musical understanding, communicative ability and technical proficiency. Results generally suggested that, besides quantity, the content and quality of practice must be examined as determinants of musical skill.

Other factors of importance for performance achievement may relate to motivation (achievement motivation, intrinsic vs extrinsic motivation), self-perception of musical competence (self-efficacy), beliefs about the causes of success and failure, and several other factors. These factors are mostly discussed and investigated regarding young performers in attempts to understand why some children/young performers sustain and even enjoy practising music while others are unable to do so. For instance, students who display adaptive mastery-oriented behaviour patterns when confronted with a problem-solving task appear to enjoy exerting effort in the pursuit of task mastery, while those displaying maladaptive, helpless patterns tend to avoid further challenges and perhaps give up. It may also be that practice results in opportunities to demonstrate one’s competence, thus contributing to strengthening the self-concept and sense of identity, which in turn generates further motivation to practise. For these and related matters, see Hallam (1998), Harnischmacher (1997), McPherson and Cormick (1999), O’Neill (1997, 1999, 2002), and O’Neill and McPherson (2002).

To my knowledge, Hallam (1997c) has provided the most encompassing review so far of all the factors that may affect learning outcomes in music. She suggested a model including a large number of influencing and interacting factors: learner characteristics (e.g. learning styles, approaches to practice, motivation, self-esteem, personality); learning environment (e.g. teacher characteristics and interventions, home environment, parental support); task requirements (e.g. nature of the task, instrument characteristics, repertoire); practice process (task-oriented strategies, person-oriented strategies); and various learning outcomes.

To conclude this section, I advise the reader to consult the volumes edited by Jorgensen and Lehmann (1997), Parncutt and McPherson (2002), and Rink (2002). Some of the contributions in these books are referred to at different places within this review, but there is much more to be gained by reading them as a whole.
PRACTICE AND TALENT
In the papers on deliberate practice, explanations of high achievement in terms of innate musical talent are rejected. Howe et al. (1998) reviewed evidence in support of, as well as contradicting, the talent account of exceptional accomplishments (see also Sloboda, 1996). They concluded that ‘differences in early experiences, preferences, opportunities, habits, training, and practice are the real determinants of excellence’ (p. 399). Open peer commentaries to this paper were given by some 30 researchers, whose opinions varied from support for the standpoint taken by Howe et al. to complete rejection of it (‘absurd environmentalism’, p. 411), affirmation that ‘inborn talent exists’ (p. 415), and claims that exceptional performance depends on both innate and environmental factors in interaction (‘fruitless polarities’, p. 411). There were also a lot of comments on choice of definitions and research strategies. This debate will certainly be ongoing.

Sight reading
Sight reading means performing from a score without any preceding practice on the instrument of that score, to perform a prima vista. Sight reading involves a combination of reading and motor behaviour, that is, to read note patterns coming up in the score while performing others just read. Earlier studies of sight reading investigated sight readers’ perception of patterns in the score (‘chunking’), the eye–hand span, eye movements and the processes involved in sight reading compared to processes in memorization of music (G99: 509–13).

Lehmann and Ericsson (1993) asked 16 expert pianists to accompany two pre-recorded flute parts from a music score. Pianists specializing in accompanying performance performed better (number of correct notes played) than pianists specializing in solo performance. Both groups started piano performance at about the same age (mean = 6.5 years), had about the same accumulated number of hours of piano practice during their life time (average 12,245 hours), and started regular sight reading at about the same age (11.2 years). However, accompanyists reported more accompanying experience, a larger repertoire of accompanying music and a larger number of public performances involving accompanying than solo performers (Lehmann and Ericsson, 1996). These results, then, indicated the importance of specialized training and relevant knowledge.

Most of the recent studies focused on the perceptual part of sight reading, that is, music reading (without performance). Typically, participants were presented with pairs of notations that displayed intervals, chords, or sequences of varying length, coherence and complexity, and had to tell whether the two stimuli in the pair were the same or different. Participants were most accurate at judging ‘same’ intervals when pairs were visually similar and least accurate at judging ‘same’ intervals when pairs were spatially
dissimilar (Gillman et al., 2002). Experts responded faster than novices; response was slower to temporally randomized material than to temporally coherent material (Waters and Underwood, 1998; Waters et al., 1997). Experts read in larger units than poorer readers (Polanka, 1995), used larger units to compare the stimuli and processed these units with fewer fixations and in less time (Waters et al., 1997). Pianists reading original and simplified versions of a Beethoven piano piece showed a larger number of fixations for the original than for the simplified versions, especially regarding regressive saccades (Servant and Baccino, 1999).

Waters et al. (1998) compared sight readers of different skill on a number of tasks. Rapid recognition of patterns (groups of notes) was the most critical factor and accounted for most of the variance in a (hierarchical) regression analysis. However, significant increases in the amount of variance explained occurred when predictors also included performance on a visual-auditory matching task (subjects should decide whether an auditory stimulus was the same as a notation they had looked at, or not) and on a kind of priming task concerning harmonies (major/minor chords). They concluded therefore that auditory skills and prediction skills (priming) contribute to skilled sight-reading ability over and above basic pattern recognition.

Eye–hand span was investigated in a couple of papers. Truitt et al. (1997) used a moving window displaying either two, four or six beats of short pieces from Mikrokosmos Vol. 1 by Bartók; only one of the staves was shown and eight participants played with one hand only on a Yamaha keyboard. The average eye–hand span for the more skilled performers was only about two beats (quarter notes; there were only quarter notes and half notes in the examples) and barely one beat for the less skilled performers. The authors suggested that an eye–hand span up to six to seven notes found in earlier studies is an over-estimate and includes notes being guessed rather than actually seen. Furneaux and Land (1999) also had eight pianists of varying skill perform short pieces, displaying both staves and playing with both hands. The eye–hand span was approximately four notes for the professional pianists and two notes for less skilled performers. They further measured a time-index, the length of time between fixation of a certain note and performance of this note; this measure did not differ between pianists of different skills. Taken together, these results indicated that the professional pianists were more efficient than less skilled pianists in chunking several notes together.

Comparison between these two reports and the reports on eye–hand span reviewed in G99 reveals several differences regarding definitions of concepts, choice of musical material, and measurement techniques which are in need of discussion before future research. For a more complete review of sight-reading and pedagogical implications, see Lehmann and McArthur (2002).
Improvisation

There were few studies on improvisation reviewed in G99 (pp. 513–15). The chapter on improvisation by Pressing (1988) remains basic reading for anyone interested in this area. In a later paper, Pressing (1998) pursued the idea of improvisation as a system of expertise, relating to standard expertise theory with its emphasis on deliberate practice and development of domain-specific skills, in this case skills such as real-time perceptual coding of events, optimal attention allocation, decision making, prediction of the action of others, error correction, movement control and others, moreover the ability to ‘integrate these processes into an optimally seamless set of musical statements that reflect both a personal perspective on musical organization and a capacity to affect listeners’ (p. 51).

Opinions differ widely regarding the proper definition and meaning of improvisation. For discussion of this and to get a broader perspective on improvisation world-wide, it is instructive to read the different contributions in the volume edited by Nettl and Russell (1998), especially the introductory chapters by Anderson Sutton (1998), Blum (1998), Nettl (1998) and Pressing (1998). Other relevant sources are Bailey (1992), Kenny and Gellrich (2002), and Sawyer (1999); these also deal with pedagogical questions.

Most of the recent empirical work concerns improvisation in jazz. Järvinen (1995) studied and found tonal hierarchies in jazz (bebop) improvisation similar to those in European art music. Järvinen and Toiviainen (2000) found that, on the whole, the basic metrical structures in bebop improvisation conform well to the traditional view of metrical organization. On strong beats, musicians tended to play important tones of the underlying tonality, especially the tonic; on other beats, the distribution of different tones was more even. Schögler (1999–2000) reported a pilot study of how visually separated musicians in jazz duets manage to achieve temporal coordination at qualitative changes in improvised music, for example, in connection with a change in meter or dynamics.

Toiviainen (1995) adopted an artificial neural network (ANN) approach to model bebop-style jazz improvisation, arguing that, in learning to improvise, the student essentially mimics the playing of other musicians, a kind of imitative learning that can be modelled and simulated by ANN. Johnson-Laird (2002), developing an earlier approach (Johnson-Laird, 1991) also dealt with bebop-style improvisation and proposed that a principle of algorithmic demands governs improvisation. There is a kind of division of labour between the creation of the underlying chord sequences and the creation of melodic improvisations. Creation of chord sequences occurs over multiple stages of creation and critical evaluation, therefore demanding working memory (or a notation) of intermediate results, whereas creation of melodies that fit the chord sequences occurs in a real-time single step and does not require any memory of intermediate results. Most jazz musicians improvise melodies to composed chord sequences, not chord sequences to composed
melodies. This approach was regarded as a computational theory of creativity. Three possible algorithms for creativity were discussed, a neo-Darwinian, a neo-Lamarckian (the one considered most likely for melodic jazz improvisation) and a compromise between the two (appropriate for the creation of chord sequences). A computer program for generating bass lines, given a chord sequence represented in symbols as input, was described and exemplified.

The role of motor processes should not be forgotten. Baily (1990, 1991) stressed the role of motor processes in planning and performance; the spatial properties of an instrument in combination with convenient motor patterns may be decisive factors for the shaping of a piece of music. This seems particularly pertinent to improvised music, exemplified in parts of *Ways of the Hand* by Sudnow which has now appeared in a rewritten account (Sudnow, 2001).

Among the recent empirical work on jazz, perhaps the most encompassing view was adopted by Reinholdsson (1998). He studied musical and social interaction in two small jazz groups, a quintet and a trio, during rehearsals and performances and developed verbal–graphical representations to describe the musical and social interaction between players. The quintet players managed to negotiate and handle unanticipated situations and moments in a flexible manner with cooperative interaction, mutual respect for, and confidence in, each other. In contrast, the trio players reported anxiety and uncertainty about how to approach problematic situations; feelings of frustration, dissatisfaction and distrust were present. The empirical work was discussed within a broad interdisciplinary framework including ethnomusicology, sociomusicology, jazz analysis and cultural sociology.

**Relations between different types of performances**

McPherson et al. (1997) used path analysis to model the relationship between five types of musical performance: performing rehearsed music, sight reading, playing from memory, playing by ear and improvising. Data were derived from the performance achievements of 101 high-school clarinet and trumpet instrumentalists. Factor analysis of possible influencing factors, explored through questionnaires, indicated four factors: Early exposure, Enriching activities, Length of study and Quality of study. Among the most important results were that performing rehearsed music was most influenced by length of study and ability to sight read, whereas the ability to improvise was most influenced by the ability to play by ear. Playing music by ear seemed to exert a positive influence on the ability to sight read. Implications for music education were discussed; see also McPherson (1995–6).

**Feedback**

Feedback in performance may be auditory, visual, tactile, kinaesthetic and perhaps vestibular; for examples and earlier reports, see G99 (pp. 515–16).
Recently, other studies have appeared. Pianists played pieces by Bach (Finney, 1997) or Chopin (Repp, 1998e) with and without auditory feedback. In both studies, there were practically no effects on various performance variables of the non-feedback condition, and listeners found it hard to distinguish between normal performances and performances without feedback. In Repp’s study, pedalling was affected: pianists changed pedals less often when they could not hear themselves. On the whole, the results implied that the motor activities involved in performance are primarily guided by the performer’s mental representation of the piece and do not depend on the existence of auditory feedback. This conclusion was also reached in a related study (Repp, 1999f), which demonstrated that pianists could generate the expressive timing pattern in the absence of auditory and kinaesthetic (piano keyboard) feedback. In attempted metronomic performance, this timing pattern was reduced but was not dependent on auditory feedback; however, the pattern was much reduced or absent when kinaesthetic feedback from the piano keyboard was eliminated (subjects tapped on a response key), thus indicating the influence of kinaesthetic feedback in this condition.

Finney’s (1997) study also included several conditions with delayed auditory feedback (DAF, 250 ms), changed pitches in feedback, and combinations of delay and changed pitch. Briefly, DAF led to performance impairment, but pitch alternation did not significantly impair performance. Pfordresher and Palmer (2002) used different amounts of DAF and found some evidence for less disruption when delayed feedback onsets coincided with binary subdivisions of the produced IOIs; deliberate attempts to count subdivisions reduced disruption at the longer delays used. Finney and Palmer (in press) manipulated auditory feedback, presence or absence, while pianists learned musical pieces and when they later played the pieces from memory. Auditory feedback during learning improved later recall, but auditory feedback at the later test did not affect recall. This result goes against results of many studies on verbal and motor learning which indicate that performance at test is best when conditions are the same at both learning and test.

**Motor processes**


The demands on motor functioning in music performance are large, indeed: there seems to be no other human activity which demands so much of time-ordered fine motoric facilities as the mastery of a musical instrument (Wagner, 1987). An obvious example is fingering in piano performance. A number of reports have resulted from a collaborative project on pianists’ fingering strategies. An interview study (Clarke et al., 1997) revealed both many common and idiosyncratic views on, for example, the use of standard
fingering, the relation between interpretation and fingering, and the teacher’s role. Parncutt et al. (1997) developed an ergonomic model of right-hand fingering in melodic fragments that involved a number of hypothetical rules, each related to a specific ergonomic difficulty, such as size of spans between finger pairs, use of weak fingers, and playing on black and white keys. Predictions from the model were tested against 28 pianists’ indication of fingering in some Czerny studies. Most proposed fingerings agreed with fingerings recommended in the model as the least difficult, but there were also many predicted fingerings that were not recommended by the pianists. Sloboda et al. (1998) compared the fingerings of masters, experts and novices. Performance accuracy and fingering consistency showed positive correlation with expertise, but there was no simple decrease in the difficulty of fingerings chosen with increased expertise. Experts and masters were more prepared than novices to squash fingers close together and to use finger patterns involving the third, fourth and fifth fingers. A study (Parncutt et al., 1999) of pianists’ performance of passages where the right and left hands played in parallel at an octave’s distance indicated that, in this situation, pianists focused more attention on the right hand than on the left hand. Some modifications of the model were proposed by Jacobs (2001); he suggested that different music-historical periods might require different fingering models owing to differences in keyboard construction at different times.

Musicians’ polymetric performance was studied by Grieshaber and Carlsen (1996). Greatest consistency in performance occurred for the faster part of the two meters, and for performance of 2:3, then for 3:4, and then for 4:5. Krampe (1997) studied age-related changes in simple motor speed, bimanual coordination and dynamic variation in expressive performance. Only bi-manual coordination showed negative age-effects and then mainly for older amateur musicians; the effect was minimal for a group of older, still active professional musicians.

Friberg and Sundberg (1999) continued earlier research (e.g. Kronman and Sundberg, 1987) on the relation between human locomotion and ritards in music. The mean body velocity of professional dancers during stopping of running (deceleration) showed a striking similarity with the mean tempo pattern of final ritards in music performances. Friberg et al. (2000) transformed vertical force patterns in walking and in dancing into sound level envelopes of a 196 Hz tone, which was repeated four times in different tempos. Listeners, spontaneously and in adjective ratings, described these stimuli in terms of various types of movements. For a general review of the relation between musical motion and performance, see Shove and Repp (1995), and Clarke (2001).

EXPRESSION MOVEMENTS
Since vision usually is the dominant perceptual sense, performers’ body movements are very important for listeners/spectators’ impression of the music and the performance. Continuing her earlier studies (e.g. Davidson,
of expressive movements. Davidson (2001) analysed body movements in singing performances with a view to studying how gestural elements help to make a performance meaningful. For example, movements during performance can communicate expressive intentions, communicate information to co-performers or to the audience, present information about the performer’s own personality, and sometimes are made to show off to the audience. Or, relating to a classification of conversation gestures, gestures in performance could be described as adaptive gestures assisting self-stimulation and being the expression of inner mental states; regulatory gestures, assisting co-ordination with co-performer(s); and illustrative and emblematic gestures assisting expression of the narrative to the audience. These categories were used in detailed analysis of movements during singing performances by Davidson herself and by the pop star Annie Lennox.

Clarke and Davidson (1998) studied a pianist’s head gestures and their relation to the musical structure and expressive intention in two performances of the Piano Prelude in E minor op. 28, no. 4 by Chopin (cf. also Clarke, 1995). Four distinct types of head movement were identified, also related to body sway, and the appearance of them at various locations in the two performances was discussed in relation to the different interpretations of the prelude: one concerned with the binary motivic structure, the other with a goal-directed, unified interpretation. Rather than suggesting any kind of causal link between conception and movement or conversely, the authors preferred to regard the performer’s conception of the musical structure and the body movements as co-determining features of the performance.

Williamon and Davidson (2002) explored communication between pianists in duet performances during rehearsal and at a recital. The pianists successively developed a set of coordinated non-verbal gestures and eye contact, especially in connection with locations in the music that were considered important for coordination and for communicating musical ideas. Talking was almost absent, over 90 percent of the rehearsal time was spent playing, described by the pianists themselves as ‘we responded to each other’s playing.’ Eye contact was critical, the performers synchronized their glances at important points in the music. Swaying movements were allied to phrase structure and tempo but also illustrative of the emotional intention of the performers.

Wanderley (1999) analysed video recordings of clarinet players’ movements in performance, such as change of posture at the beginning and during phrases; slow continuous gestures in upward direction during long sustained tones; and fast sweeping movements of the bell accompanying short staccato notes. Some movements may have been due to particular technical difficulties, whereas others seemed intended to express information beyond that conveyed by the sound. In another study (Wanderley, 2002), clarinet players wore infrared (IR) markers, which were tracked by three IR cameras, allowing study of their movements in a standard performance, an expressive performance and a performance in which the player tried not to
move the instrument at all. There was a strong correlation between the same player’s movements at the same points in the score in repeated performances, indicating that the gestures were an integral part of the performance. Different clarinetists showed different overall patterns; however, some movement features related to structural characteristics were largely unvaried across performers.

A broader review of the different functions of body movements in performance is given by Davidson and Correia (2002); see also the works of Reinholdsson (1998) and Schögl (1999–2000) mentioned under the Improvisation section. To sum up, this survey indicates an increased interest in expressive movements and their function in music performance.

**CONDUCTORS’ EXPRESSIVE GESTURES**

Expressive gestures used by classical orchestra conductors were described by Boyes Braem and Bräm (2000). It seems that many of the expressive gestures of conductors are related to gestures accompanying speech and to sign language of the Deaf (e.g. the set of hand shapes used by conductors includes those found in most sign languages around the world). It was claimed that the expressive gestures are based on metaphoric/metonymic connections between aspects of the music and physical experiences in human beings’ everyday life, such as handling objects (e.g. grasping, hitting, letting go, supporting, pointing) and biological functions of the body (e.g. hearing, smelling). Because they have so much in common with other aspects of human experience and communication, conductors’ expressive gestures can function effectively, without further verbal explanation, with musicians from many different cultures (however, there are culturally specific gestures as well). An instructive appendix displayed drawings of many conducting gestures – most of them easily recognizable by readers with some experience of watching conductors – described in terms of metaphoric connection, intended musical effect, and the gestures themselves (hand shapes, movements, location).

While dealing with conductors, it may be interesting to note that, in comparison with pianists and non-musicians, experienced conductors demonstrated better auditory localization in the peripheral space, a factor of importance in their auditory supervising of an orchestra (Muente et al., 2001).

**Physical factors**

Music performance puts very high demands on the body and there is a high frequency of musicians having medical problems associated with their performance: pain and tenderness in muscles, problems with shoulders, neck, back, fingers, hands, arms, palsies, hearing impairment, neurological disorders, high stress levels and others (G99: 557–61). Numerous reports on
such problems continue to appear; see, for example, in the journal *Medical Problems of Performing Artists*. I have not been able to follow all of this literature. Here are some recent works concerning hearing problems.

Chesky and Henoch (2000) found that about 20 percent of a sample of more than 3000 musicians reported problems with hearing. The highest rate of occurrence was in rock and jazz musicians and in musicians who played amplified instruments, drumset and brass instruments. Sound levels often exceeded recommended maximum levels to avoid hearing impairment, in the orchestra pit in Broadway shows (Babin, 1999) and during rock/jazz performances (Kähäri, 2002). A study of 140 classical and 139 rock/jazz musicians (Kähäri, 2002) showed slightly worse hearing thresholds in rock/jazz musicians than in classical musicians; among the latter, brass players had somewhat worse hearing thresholds. In rock/jazz musicians hearing loss, tinnitus and hyperacusis were significantly more frequent than in reference populations. Sataloff (1997) discussed hearing loss in singers and other musicians regarding their causes, site of damage, legal aspects and treatment. Behroozi and Luz (1997) reviewed research on noise-related ailments on performing musicians, and Lockwood et al. (2001) discussed definitions, causes, treatment and prevention of tinnitus in musicians.

At a more general level, a long-term follow-up of treatment outcome for instrumentalists (Lederman, 1995) showed that a large majority of instrumental musicians seen for various playing-related symptoms can be helped by carefully chosen treatment programmes; however, it also showed that 12 percent of the instrumentalists had given up playing entirely.

**Psychological and social factors**

This section in G99 (pp. 561–77) dealt with musical development in performers, personality of performers, music as an occupation and performance anxiety. This update includes only limited coverage of recent research in these areas.

**DEVELOPMENT**

Developmental studies reviewed in G99 (pp. 561–4) mainly concerned interview studies tracing the musical development of professional classical musicians or advanced students; also, at that time (1995), the recently launched concept of deliberate practice.

Some chapters in a recent book on musical identities (MacDonald et al., 2002) deal with the development of musical identities in children, young musicians and solo performers. On the basis of interviews with 12 families, Borthwick and Davidson (2002) found, briefly, that children’s musical identity was shaped primarily by the responses and values in the family. Families who adopted music as a core value allowed music to influence daily routines; music was for all members of the family to discuss, listen to and perform.
Parents expected the children to be involved in musical activities and ensured that they practised their instruments daily. In two-child families, the first-born child was considered more musically talented than the second-born, while parents with three children perceived each child as unique with their own niches rather than being compared with one another. Lamont (2002) studied and discussed how teachers and the school environment may influence the development of children’s musical identity.

Among factors that may influence the development of a solo performer, Davidson (2002) mentioned early strong experiences of music, frequent exposure to music and performance contexts, and support of key others such as peers, family and teachers; generally, both musician and performance skills develop optimally in unthreatening situations. Other crucial factors concern motivation and personality. Motivation is strongly affected by self-belief – for instance, if one considers one’s ability in a certain domain as fixed once and for all, or, in contrast, as continuously improvable. In a follow-up study, eight years later, of successful music students who all intended to become professional musicians it turned out that only half of them were professionals, while the other half had kept music as a hobby, pursuing other careers. All the professionals had received external support from teachers, parents and peers. Moreover, over time music had become the key determinant of self-concept; music making had provided positive physical and psychological benefits; and music performance had become a critical means of self-expression. Davidson also suggested that learning how to present the self on stage is a critical variable for performers, referring to her own studies in this area (Davidson, 2001; Davidson and Coimbra, 2001).

PERSONALITY
A number of different and rather scattered approaches concerning the personality of musicians were cited in G99 (pp. 564–6). Fortunately, the present discussion on this topic can be brief, simply referring the reader to Anthony Kemp’s book *The Musical Temperament: Psychology and Personality of Musicians* (1996). This work is a landmark in this area and, to the best of my knowledge, the first book ever on musicians’ personalities. In his exposition of this extremely complex area, Kemp referred to several different theories of personality, especially Cattell’s trait theory, the typologies of Eysenck, Jung and the related typology of Myers-Briggs, and the so-called ‘big five’ dimensions of personality proposed by Costa and McCrae. His discussion of musicians’ personalities focused on four dimensions: introversion, independence, sensitivity and anxiety, and the characteristics of different groups of musicians – orchestral musicians, keyboard players, singers, conductors, popular musicians, composers and music teachers – were discussed using these dimensions, and other related material, as a framework. Anyone interested in what kind of people musicians are should read this book and reflect on it in relation to one’s own opinions and experiences.
As Kemp’s own research has concentrated on musicians in the classical genres, it has been suggested that there may be a certain bias in that direction in his account. For instance, in a recent study of the personality of rock musicians, Gillespie and Myors (2000), relying on the ‘big five’ dimensions (neuroticism, extraversion, openness, agreeableness, conscientiousness) found that rock musicians seem to share a common core of traits: higher than average in neuroticism and openness, average in extraversion, and below average in agreeableness and conscientiousness. While it may be a pretty popular view that rock musicians are extraverted, the results indicated that ‘rock musicians, while not showing the introversion of classical instrumentalists, do not seem to be especially extraverted either’ (p. 161).

MUSIC AS OCCUPATION
Earlier studies have indicated that working as a musician may lead to many problems: irregular working times; long periods away from home on tours; difficulties in sleeping after intense concentration during concerts; tension among ensemble members; fear of competition, with accompanying feelings of self-doubt and depression; little or no influence on decisions concerning work; incompetent conductors; powerlessness in relation to managers; high demands on technical perfection; performance anxiety and so on (1999: 566–9). Such problems were also frequently mentioned in recent studies in relation to such different categories as popular musicians (Raeburn, 1999, 2000) and opera singers (Sandgren, 2002); also, in a comparison between dancers and musicians (Hamilton et al., 1995), occupational stress was most severe for the dancers.

Connecting to her earlier papers (Raeburn 1987a, 1987b) on occupational stress in professional rock musicians, Raeburn (1999, 2000) pointed out that the types of problems popular musicians meet – such as those just mentioned – should also be considered in relation to the common image of the popular musician, especially the rock musician, as a deviant or outsider, expected to show emotional extremes and behaviour that challenge conventional societal norms. Moreover, the demands of the music industry contribute to create financial and emotional uncertainty for popular musicians. Raeburn offered a general review of psychological issues and treatment strategies in popular musicians with special focus on depression, anxiety, and substance abuse and dependence.

In a study of opera singers, in addition to the problems already mentioned, Sandgren (2002) identified some specific problems for opera singers: preoccupation with the risk of vocal indisposition, leading to excessive use of vitamins and herbal products, and avoidance of places thought to lead to risk of infections; worries about significant others’ opinions about their performance; much self-criticism; and the need to continually test the presence and quality of the voice. Positive correlations were found between worry about others’ opinions and variables related to somatic problems, depression and
performance anxiety. Some gender differences appeared. For instance, for men, strong emotional singing experiences were associated with a feeling of complete mastery of the voice – that the technique was obeying – whereas women stressed a feeling of total presence and expression of emotions to a responsive audience.

With regard to workplace conditions for classical orchestral musicians in the UK and Germany (Harper, 2002), a questionnaire revealed that the UK sample \((n = 817)\) viewed noise and cramped playing positions as the issues that caused greatest concern; for the German sample \((n = 187)\), they were noise and ventilation. Hearing protectors inserted into the ear were common but not judged particularly successful.

Despite all these problems, most musicians stay in their profession and would choose it again. Making music has many positive aspects – playing together with people of like minds, performing to responsive audiences, meeting with famous musicians and, above all, the elusive intrinsic rewards offered by close contact with this phenomenon called music (G99: 567–8). Davidson (1997) reviewed and discussed a number of relevant topics within the social psychology of music performance: historical performance practices; performance etiquettes (e.g. expected dress and behaviour); presence and appraisal by key individuals and others; social dynamics in groups, exemplified by studies on string quartets and orchestras regarding leadership, individual members’ roles, group cohesion or dissolution, handling of conflicts, etc; communication between performers, verbal and non-verbal (eye contact, facial expression, various body movements); and expressive body movements accompanying performance – see also Davidson (2001) in the Motor processes section. Some of these points were explored in a study (Davidson and Good, 2002) of rehearsals and concert performance of a newly formed string quartet with a focus on the social interaction among the members and the efforts to coordinate their performance at different places in the musical structure, verbally and by gestural markings of dynamics, entrances and exits. Similar processes also appeared in Reinholdsson’s (1998) study of interaction in small jazz groups, see the Improvisation section.

PERFORMANCE ANXIETY
Prevalence, symptoms, contributing factors and coping strategies concerning performance anxiety were treated in some detail in G99 (pp. 569–77). Further reviews have appeared in Brodsky (1996), Steptoe (2001), Wilson (2002), and Wilson and Roland (2002) to which the interested reader is referred; here only a few recent papers are mentioned.

With regard to the medical treatment of performance anxiety, Lederman (1999) suggested the use of a low dose of a beta-blocker to prevent debilitating physiological reactions (e.g. sweating, tremor) but only after prescription by a physician. Sataloff et al. (1999) discussed some situations in which medical treatment of performance anxiety may be appropriate but emphasized
that performers who require alcohol or another medication routinely to perform their daily activities in their profession should not have their symptoms medicated but be referred to a psychological professional with expertise in arts medicine.

Prevalence of music performance anxiety was studied for Dutch orchestral musicians (Van Kemenade et al., 1995) and for Norwegian music students (Kaspersen and Goetestam, 2002). Ryan (1998) found evidence of performance anxiety already in young piano students (aged 12 years), and so did Hallam (1997a: 102) for even younger string players. Young musicians perceived performance anxiety as an entity detrimental to performance (O’Neill, 2002), occupying a position of power relative to the individual with the capacity to control their actions, an ‘enemy’ that has to be fought; they expected to learn how to do this better as they gained more experience of performance.

Leblanc et al. (1997) found increased performance anxiety with a larger audience. Esplen and Hodnett (1999) found reduced performance anxiety following a guided imagery intervention. Rife et al. (2000) found that lower levels of performance anxiety were associated with higher levels of cognitive flexibility. Osborne and Franklin (2002) examined the relation between music performance anxiety and social phobia. They concluded that social phobic models may provide a valid account of the cognitive processes occurring in music performance anxiety; however, some results suggested that high music performance anxiety is not synonymous with social phobia.

The impression from reading many reports on performance anxiety is that there are often flaws and limitations, sometimes admitted, sometimes left without comments. Brodsky (1996) provided a highly relevant, critical review of current research practices on music performance anxiety and concluded that diversity of conceptions concerning stress and anxiety, inadequate sampling, inept screening criteria, and unreliable and invalid assessment procedures may ‘have led performing arts medicine practitioners and researchers down the wrong avenues’ (p. 96).

**Evaluation of performance**

Evaluation of performances occurs in the everyday activity of music critics, music teachers and musicians. However, there are hardly any agreed criteria either for what should be judged, or for how the judgements should be made. Judges may be unaware of what criteria they actually use in their assessments. Most earlier studies were concerned with evaluation of performances by music students at different levels of skill. Reported interjudge reliabilities varied considerably (G99: 577–9).

Three recent reports dealt with evaluation of performance by music students. Saunders and Holahan (1997) arranged for solo performances by 926 students on woodwind and brass instruments to be assessed by 36 judges regarding tone, intonation, technique/articulation, melodic accuracy,
rhythmic accuracy, tempo and interpretation. Reliabilities were considered sufficient and, in a regression analysis, the total score was predicted by a weighted combination of tone, technique/articulation, rhythmic accuracy, interpretation and achievement in a sight-reading performance. Bergee (1997) used similar scales for assessment of undergraduate music student performances regarding voice, percussion, brass, woodwind and stringed instruments. Judges were faculty members and students (peers), including the performers themselves (looking at video recordings of the performances). Most reliabilities were acceptable. Correlations between faculty and peer evaluations were high, but self-evaluation correlated poorly with faculty and peer evaluations. Daniel (2001) discussed the desirability of students’ self-assessment of their performance. He asked students to watch video recordings of their own performance and provide self-critical reports on them. Almost half of the students were very critical of their own playing (cf. the results of self-evaluation in Bergee’s report cited earlier), but most students also regarded this type of self-assessment as very valuable for enhancing their performance. This type of self-assessment may help students to develop skills in thinking independently and reflecting on their practice and performance; in other words, to learn appropriate strategies for practice and performance (cf. Hallam, 1997a, 1997c; McPherson and McCormick, 1999; Nielsen, 1999b, 2001: see the Practice section).

Criteria used in assessment of student performance are dominated by various technical aspects, such as intonation, articulation, tempo, melodic and rhythmic accuracy; interpretation may be included but seems to play a minor role. At successively higher levels of performance skills, other features come to the fore. Two recent studies aimed to explore the types of criteria used in the assessment of singers’ live performance (Davidson and Coimbra, 2001) and expert pianists’ recorded performances of a Chopin Etude (Thompson et al., 1998). Both studies basically used a qualitative approach: questionnaires and interviews with judges (Davidson and Coimbra) and repertory grid analysis following judges’ proposals of bipolar constructs appropriate for describing performance expression (Thompson et al.); the degree of formalization was considerably higher in the latter study. Analysis of judges’ comments on singers’ performances indicated a number of criteria for assessment: choice of appropriate repertoire; technical control (e.g. vocal support, diction, integration of different registers in the voice; surprisingly, nobody mentioned the timbre of the voice); the physical appearance/appeal of the singer; the singer’s bodily communication through facial expression and eye contact; artistry, that is, communicating meaning to the audience, seeming committed to interacting with the audience; the singer’s presence (focused on transmitting musical intentions), and ‘performing personality’. The most critical variables were singers’ physical appearance allied to their control of the voice in order to project emotional expression.

Physical appearance was of course not relevant in the assessment of recorded piano performances. The repertory grid technique resulted in a total
of 14 constructs (criteria), for instance, right-hand expression, phrasing, dynamics, rubato, tonal balance, pedalling, articulation, tempo and expression in certain bars. Constructs that showed the highest correlations with ratings of overall preference were right-hand expression, phrasing and expression at the end of the piece. Although individual judges differed in regard to the constructs used and the ratings on them, there was an overall agreement concerning the ranking of the performances.

Judgement of performances was also conducted in many of the reports dealing with measurements of performances. In some cases, judgement referred to relatively simple perceptual or emotional variables, as in many studies on emotional expression in performance; in other cases, judgements concerned more subtle aspects. For instance, commenting upon experienced musicians’ ratings of the approximate 100 different performances of Chopin’s Etude in E Major, Repp (1999b) discussed many different possible criteria that may have been used but found no convincing evidence for any of them. His personal guess was that the elusive quality called ‘tone’ or ‘touch’ in piano performance was the most crucial aspect, but its precise meaning as well as its correlates in physical variables are not well understood, to say the least. Perhaps such qualities defy common verbal description and can only be hinted at by means of metaphors or by some kind of non-verbal technique. In another paper, Repp (1998b) suggested that musical structure has kinematic implications and that listeners in a way ‘move as they listen’ (p. 809), an idea for which I have much sympathy and that might perhaps be explored by means of listeners’ expression of perceived motion characters on something like Clynes’s (1977) sentograph.

In conclusion, whatever level of performance skills and whatever kind of music performed, much work remains to establish adequate criteria for the evaluation of music performance.

Concluding remarks

Although not all published reports have been included in this review, it is obvious that music performance research is in a very active stage. There were many new reports in all the different areas of performance research displayed in Table 1. Measurement of performances is still the largest area in terms of the number of reports, and results of measurements are now increasingly incorporated and interpreted in different models aimed at finding general principles behind the mass of data. In comparison with the distribution of reports in G99 (Figure 1), the most noticeable change is the large increase of reports regarding the planning of performance, especially regarding practice before performance. This also means that results from performance research are brought closer to matters of music teaching and learning (Jorgensen and Lehmann, 1997; Parnicutt and McPherson, 2002; Rink, 2002). The number of contributions to the planning of performance would probably be still more
increased if the recent attempts at bringing music analysis and empirical performance research closer to each other – as seen, for example, in the two volumes edited by Rink (1995, 2002) – had been included. I was unable to do so in the present review. I was also unable to include contributions in music aesthetics bordering on empirical research as, for example, in the recent book by Davies (2001). It was frustrating to see that, although Davies and I (G99) both had hundreds of references to papers on performance, we did not have a single reference in common. Generally, it is high time that music psychology came into closer contact with neighbouring and overlapping disciplines, such as music analysis (already mentioned), music anthropology, music aesthetics (Gabrielsson, in press b), music philosophy, etc. All of them deal with different aspects of the same phenomenon, that is, human beings’ relationship to music.

Finally, I want to stress some points that have been partly remarked upon earlier. Music performance should be studied as much as possible: (a) in musically relevant contexts to ensure ecological validity; (b) in relation to performers’ intentions, and listeners’ experiences and reactions; (c) both as process and product; and (d) considering that performances are (should be) aesthetic objects. We are still far from understanding the aesthetic aspects of music performance and experience. To come closer, we should not only strive to find broad ‘lawful’ relations in music performance but also study the characteristics of single impressive performances – in other words, adopt both nomothetic and ideographic approaches. Eric Clarke (1995: 52) made some very pertinent remarks regarding this question.

In a review of Deutsch (1999), Repp (1999g: 274) found that the chapter on performance (=G99) left the impression that ‘music performance is exceedingly complex’. I am afraid that this impression remains even after the present review. Music performance is a very complex matter – also evident in Repp’s own papers on performance, by the way – but it is a complexity that apparently continues to constitute a challenge and inspiration to many of us, musicians as well as researchers.

ACKNOWLEDGEMENTS

I express my sincere thanks to many colleagues over the world who have generously sent me reprints and other material and to Susan Hallam for advice on the English.

REFERENCES


Gabrielsson, A. (in press b) ‘Reflektioner kring förhållandet mellan musikestetik och musikpsykologi (Reflections on the relation between music aesthetics and music psychology)’, *Nordisk estetisk tidskrift*.


ALF GABRIELSSON held a professorship in the Department of Psychology at Uppsala University (retired in 2001) where he taught general psychology and music psychology, as well as music psychology at the Royal University College of Music in Stockholm. He is a member of the Royal Swedish Academy of Music, was President of the European Society for the Cognitive Sciences of Music (ESCOM) 1997–2000 and was awarded ESCOM Honorary Life Membership in 2002. He has published numerous papers on musical rhythm, music performance, music experience (especially strong experiences related to music), and on music psychology in general, also papers on various topics in acoustics and audiology.

Address: Department of Psychology, Uppsala University, Box 1225, SE-751 42 Uppsala, Sweden. [email: alf.gabrielsson@psyk.uu.se]